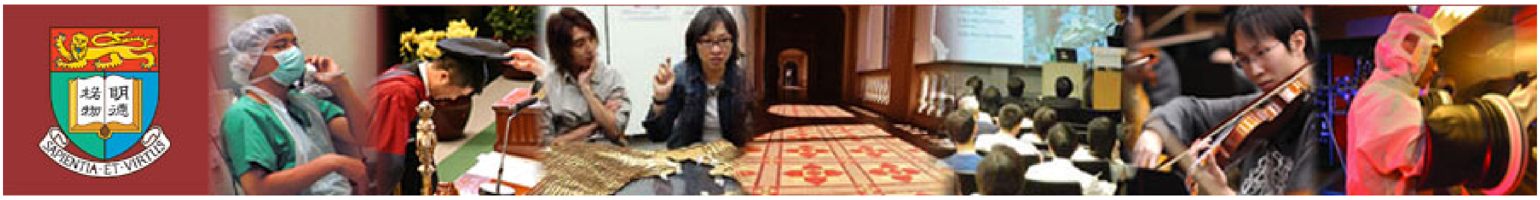


Title	**Spectral and temporal features of tense-lax vowel contrast produced by Cantonese speakers of English: a comparative study**
Other Contributor(s)	**University of Hong Kong.**
Author(s)	**Lam, Sin-ting, Stephanie**
Citation	
Issued Date	**2007**
URL	**http://hdl.handle.net/10722/55511**
Rights	**Creative Commons: Attribution 3.0 Hong Kong License**

Running Head: Spectral and Temporal Features of Tense-Lax Vowel Contrast produced by Cantonese Speakers of English – A Comparative Study

Spectral and Temporal Features of Tense-Lax Vowel Contrast produced by Cantonese Speakers of English – A Comparative Study

Lam Sin Ting Stephanie

The University of Hong Kong

A dissertation submitted in partial fulfillment of the requirements for the Bachelor of Science (Speech and Hearing Sciences), The University of Hong Kong, June 30, 2007.

Abstract

The spectral and temporal features of tense-lax vowels (/i-ɪ/, /æ–ɛ/, /u-ʊ/) produced by Cantonese speakers of English were examined. The F1 and F2, Euclidean distances, the durational differences and the duration of vowels produced in the /hVd/ context by 40 Cantonese speakers of English were measured. Perceptual discrimination tasks were performed. Results indicate that the /i-ɪ/ vowel pair was significantly different in all acoustic features, while /u-ʊ/ showed significantly different in duration and F2, and /æ–ɛ/ exhibited significantly different F1 values. Cantonese speakers produced shorter ED and DD in all vowel pairs except /i-ɪ/. Perceptual tasks also reflected less distinguishable tense lax contrast produced by Cantonese speakers of English. The present study indicated that Cantonese speakers might make use of durational cues for producing tense/lax vowel contrast. The lack of tense-lax vowel contrasts and the presence of short/ long allophones in Cantonese may influence the acquisition of tense/ lax vowels in English.

Introduction

Studies have found that learning of a second language (L2) tends to show inaccuracies in both the perception and production of L2. L2 learners tend to perceive and produce L2 phonetic segments differently from what the native speakers of L2 do (Flege, 1989; McAllister, Flege, & Piske, 2002; Chen, Robb, Gilbert, & Lerman, 2001). Issues concerning such inaccuracies in L2 production raised by previous studies included: (1) learners perceive and produce L2 more accurately if they gain more experience in L2 (Flege, Bohn, & Jang, 1997); and (2) it is suggested that the degree of accuracy in perceiving and producing L2 depends on the background of L1 (Flege et al, 1997; Jenkins, 2000).

The second issue is based on the fact that the similarity and difference between the sound systems of L1 and L2 have competing influence on the learning of L2 pronunciations. To explain this, two hypotheses have been put forth by different researchers (e.g., Eckman, Elreyes, & Iverson, 2003). The Contrastive Analysis (CA) Hypothesis assumes that there is phonological transfer from L1 to L2 (Flege et al, 1997; Jenkins, 2000). Features of L1 that are similar to those of L2 can be transferred or can facilitate the acquisition of L2. Features of L2 that are different or not present in L1 can cause difficulty in the perception and production of L2 (McAllister et al., 2002; Jenkins, 2000). The second hypothesis is related to the problem arising from similar features between L1 and L2 because phonemes with similar features do not imply allophonic equivalence. Eckman et al. (2003) illustrated three different

L2 learning situations: (1) native language does not contain neither of the sounds that are contrast in the target language; (2) native language contains only one of the contrastive sounds; and (3) native language contains both contrastive sounds in the target language but they are present as allophones. They claimed that situation 3 causes the greatest learning difficulty since it is hard to split allophones in L1 into contrastive phonemes in L2. Some studies claimed that L2 sounds similar to that in L1 are difficult to acquire than those dissimilar ones (Eckman et al., 2003).

For examples, though /i, u, ɑ/ are present in both Mandarin and English, vowels may not be acoustically identical across the two languages (Chen et al. , 2001). Learners may judge such acoustically different phonemes to be the same (Jenkins, 2000). Another example is the acquisition of French vowel [ü] by native speakers of English. The French vowel [ü] is absent in English, and the closest counterpart appears to be [u]. Flege (1987;1990) found that English speakers acquired the new French vowel [ü] more accurately than the French vowel [u]. This finding was explained by using the hypothesis that, when acquiring the French vowels, native speakers of English substituted the French vowel [u] by English [u] directly, whereas they learned French [ü] as a new and distinct sound.

Investigation on how Cantonese speakers produce English as an L2 can provide insight into the influence of phonetic features on L2 acquisition. Cantonese is a tonal language. It has over 98% consonant vowel consonant (CVC) and consonant vowel (CV) monosyllabic structure (Patrick, Wong, & Diehl, 2003). There are six contrastive tones including three entering tones in Cantonese varied mainly by fundamental frequency. As a tone language, a change in the tone level alters the meaning of the word. English is non-tonal language and a change in tone leads to a change in intent only (Wong, & Diehl, 2003).

In a recent study on Hong Kong English (HKE) phonology, Hung (2000) examined the segmental aspects of HKE by using spectrographic analysis. It was reported that the phonemic inventory (vowel and consonant systems) of HKE is simpler than that of native English. One of the significant findings on the systematic features of HKE vowel system is the lack of tense/lax distinction. Instead of having 11 vowel contrasts same as that in English, HKE has only 7 vowel contrasts. Hung (2000) suggested that the difference may be due to the influence of Cantonese which also has simpler vowel system. The presence of allophones instead of phonemic contrasts in Cantonese alters the transfer of sounds from L1 (Cantonese) to L2 (English). When learning English as an L2, Cantonese speakers undergo allophonic split in order to separate the allophones into contrastive phonemes.

Cantonese Vowel System

In Cantonese, there are 8 vowel phonemes /a, i, u, ɛ, ɔ, æ, y, ɐ/ and 13 allophones: /a, ɪ, u, ɛ, ɔ, æ, y, ɐ, i,ʊ, e, o, ɵ/. Acoustic studies revealed that duration plays an important role in the perception and production of vowels if the vowels occupy the same vowel space such as /i/ and /ɪ/, but differ in their relative duration (Bauer & Benedict, 1997). Within the thirteen allophones, seven are long vowels /a, i, u, ɛ, ɔ, æ, y/ and six are short /ɪ, ɐ, ʊ, e, o, ø/. Five pairs of long and short vowels /i, ɪ/, /ɛ, e /, /æ, ø/, /u, ʊ/, /ɔ, o/ are identified. The long and short members of each pair are allophones of each other but not regarded as tense/lax vowel phonemes (Bauer & Benedict, 1997). Members of the same allophones do not occur in the same context. For examples, [i] only occurs before labials and dentals whereas [ɪ] only occur before velars (Bauer & Benedict, 1997).

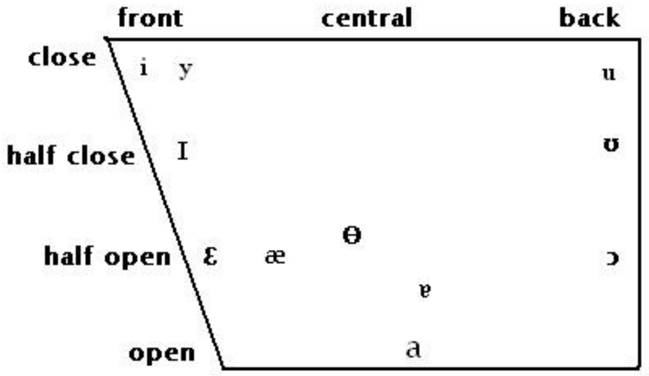


Figure 1. Vowel rectangle of Cantonese

English Vowel System

In English, there are 11 distinctive vowel phonemes /ɪ, i, e, ɛ, æ, ʌ, u, ʊ, o, ɔ, ɑ/ varying in features of tongue articulation: tongue position and height. There are 5 long vowels /i, ɛ, u, ɔ, ɑ/ and 6 short vowels /ɪ, e, æ, ʌ, ʊ, o/. They are regarded as tense-lax vowels contrast. (Olive, Greenwood, & Coleman, 1993).

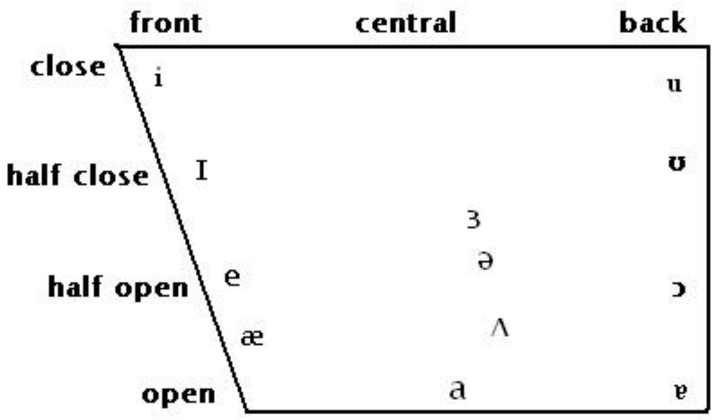


Figure 2. Vowel Rectangle of English

Purpose of the study

Research on Cantonese speakers of English has been limited. Studies of English produced by native Cantonese speakers were oversimplified and without detailed and in-depth description of HKE phonology (Hung, 2000). Early studies focused on the description of the “common error” in HKE pronunciation such as substitution of [l] for [n] or [w] for [v] (Hung, 2000). Peng and Setter (2000) carried out an in-depth investigation on the phenomenon of consonant cluster simplification in Cantonese speakers of English. Though the study successfully pointed out the systematic and unique phonological rules applying to Cantonese speakers of English (i.e. the deletion of alveolar stops in consonant cluster within specific context), generalization was limited due to the small sample size and the lack of an in-depth study focused on vowel production.

The current study attempted to investigate the differences in spectral and temporal aspects of tense/lax English vowels produced by native Cantonese and native English speakers. It was based on the study of tense/lax vowel contrasts produced by Mandarin speakers of English (Chen, 2006). Chen evaluated the articulation of tense-lax vowel pairs in Mandarin accented English (MAE) by examining the changes in temporal and spectral characteristics of the vowel pairs (Chen, 2006). Results indicated that Mandarin speakers differed significantly from the American English speakers in differentiating the English tense/lax vowels. They made use of temporal feature rather than spectral feature to indicate tense/lax contrast. The finding was explained by the absence of lax vowels Mandarin vowel systems. Mandarin speakers might fail to perceive accurate temporal and spectral difference between the tense lax vowels present in English or they could not perform the articulatory modification to indicate the contrast in production (Chen, 2006).

Different from consonant production, vowels are produced without absolute contact of articulators and vowel sounds form a continuum without distinct boundaries (Ladefoged, 1975). Subtle changes may influence the vowel quality such as the neighboring consonants, the rate of speech, the stress level of the syllable, etc.(Olive et al., 1993). Tense-lax vowel differentiation is also an aspect of vowel articulation (Chen, 2006). Tense and lax vowels can be differentiated by the vowel space (i.e. F1-F2 plot) and duration. Tense vowels are articulated at extreme position of the vocal tract leading to less centralized F1-F2 vowel space and longer duration (Chen, 2006; Ladefoged, 1975). Lax vowels are more centralized and demonstrated shorter duration (Ladefoged, 1975). Investigations of L2 vowel production by L1 speakers are important in providing insight to the acquisition of L2 phonology.

According to the Contrastive Analysis Hypothesis, it is hypothesized that the tense/lax vowel contrast is transferred from L1 to L2 if both languages have similar tense-lax vowel contrasts. However, in Cantonese, these are only allophones which differ in duration of time but not tense/ lax vowel phonemes in a way similar to the vowel system of English. The spectral and temporal aspects of tense-lax vowel pairs produced by Cantonese speakers of English will be significantly different from those produced by native English speakers with respect to both the temporal and spectral features. It is predicted that Cantonese will produce tense/lax with relatively more similar to Native English speakers than that produced by Mandarin. It might be related to the presence of more short and long allophones pairs in Cantonese and Cantonese might have better perception on temporal and spectral aspects on tense/lax vowels.

Method

Subjects

Forty adults (20 females and 20 males) who spoke Cantonese as L1 and English as L2 participated in this study. The average age of the Cantonese male speakers was 21 years (range: 19-27 years), and that of female speakers was 21 years (range: 19-23 years). The recruitment criteria included : (1) they received university education; (2) they did not major in linguistics and English nor had they attended English phonetics courses; (3) they were above 19 years old; (4) they were native Cantonese speakers; (5) they were using English as a medium of education in university; and (6) they had the ability to read English fluently and (7) they attained grade C or above in the oral part of *Use of English* in Hong Kong Advanced Level Examination (HKALE). All subjects were reported to be free of speech and language problems. They were also tested to have no hearing difficulties through hearing screening examination.

Listeners

A group of ten native American English speakers was recruited to perform the perceptual analysis. The recruitment criteria for the listeners were (1) the ability to speak and understand native American English; (2) have resided in the North America for at least 7 years. All listeners were free of speech, language or hearing impairment.

Speech materials and data collections

The speech materials included /i, ɪ, æ, ɛ, u, ʊ/ placed in the /h**V**d/ (i.e. “heed”, “hid”, “head”, “had”, “who’d”, “hood”) context produced in a carrier phrase: ‘Say /h**V**d/ again’ (Chen et al., 2001). The six vowels represented three pairs of tense-lax vowels /i-ɪ/, /æ -ɛ/, /u-ʊ/ which were present in the Cantonese vowel system as allophones. Table 1 shows the differences in tongue position and tongue height, and duration of these vowels.

Table 1. Differences in tongue position, tongue height, and duration between vowels in /i-ɪ/, /æ -ɛ/, and /u-ʊ/ (Edwards, 2003)

Vowel pair	Tongue position	Tongue height	Duration of time
/i-ɪ/	front	high	longer duration of /i/ than that of /ɪ/
/æ -ɛ/	front	mid to low	longer duration of /æ / than that of /ɛ/
/u-ʊ/	back	high	longer duration of /u/ than that of /ʊ/

A total of 720 (6 vowels x 40 speakers x 3 trials) phrases were recorded. The stimuli were presented by using PowerPoint slides in which the phrases were prepared. The order at which the stimuli were presented was randomized for each subject, in order to avoid the practice effects (Peterson & Barney, 1952). Another Microsoft PowerPoint file was prepared by entering the six vowels /i, ɪ, æ, ɛ, u, ʊ/ in the /h**V**d/ context accompanied by the corresponding sound file. The sound files were recorded from productions made by a male speaker of American English.

The recording experiment was divided into two sessions: the practice session and the reading session. Each session lasted for around 15 minutes. During the practice session, subjects were presented with the PowerPoint slides consisted of “heed”, “hid”, “head”, “had”, “who’d”, “hood” and the corresponding sound clips(Appendix A). Subjects were allowed to listen to the sound samples in order to familiarize themselves with the words they were going to produce. The sound samples could be repeated any times until the end of the practice session. Instructions on the later recording session were also provided during the practice session.

In the recording session, subjects were presented with 18 slides consisted of carrier phase "Say /hVd/ again" with vowels /i, ɪ, æ, ɛ, u, ʊ/ in a random order. Subjects were asked to produce each phrase shown in the slide. They were instructed to speak at their habitual speech tempo and pitch. A total of 18 phrases per speaker were recorded. Productions were judged by another experimenter. Incorrect production was discarded and the speaker was asked to be repeated after listening to the corresponding sound clip once.

All recordings took place in a sound isolated room in the Prince Philip Dental Hospital. Vowel samples were recorded and digitized by using a high quality digital recorder (Microtrack 24-bit/96 kHz Professional 2-Channel Mobile Digital Recorder) with a high-quality microphone of low distortion and wide frequency response (Colton & Casper, 1996). During the recording, the microphone was maintained at a constant distance of 20 cm from the speakers' mouth.

Acoustic analyses

The vowel duration, first and second formant frequencies (F1 & F2) of each vowel sample were obtained by using a speech software package (Praat: doing phonetics by computer, Version 4.4.33). To begin with the acoustic analysis, waveform and wide-band spectrogram (window length = 0.005s) for each /h**V**d/ syllable were displayed (see Figure 3). Changes in vowel duration and vocal tract vowel space were determined and calculated to differentiate between tense and lax vowels (Flege, Bohn, & Jang, 1997; Kewley-Port, Atal, 1989; Port, Rotunno, 1979).

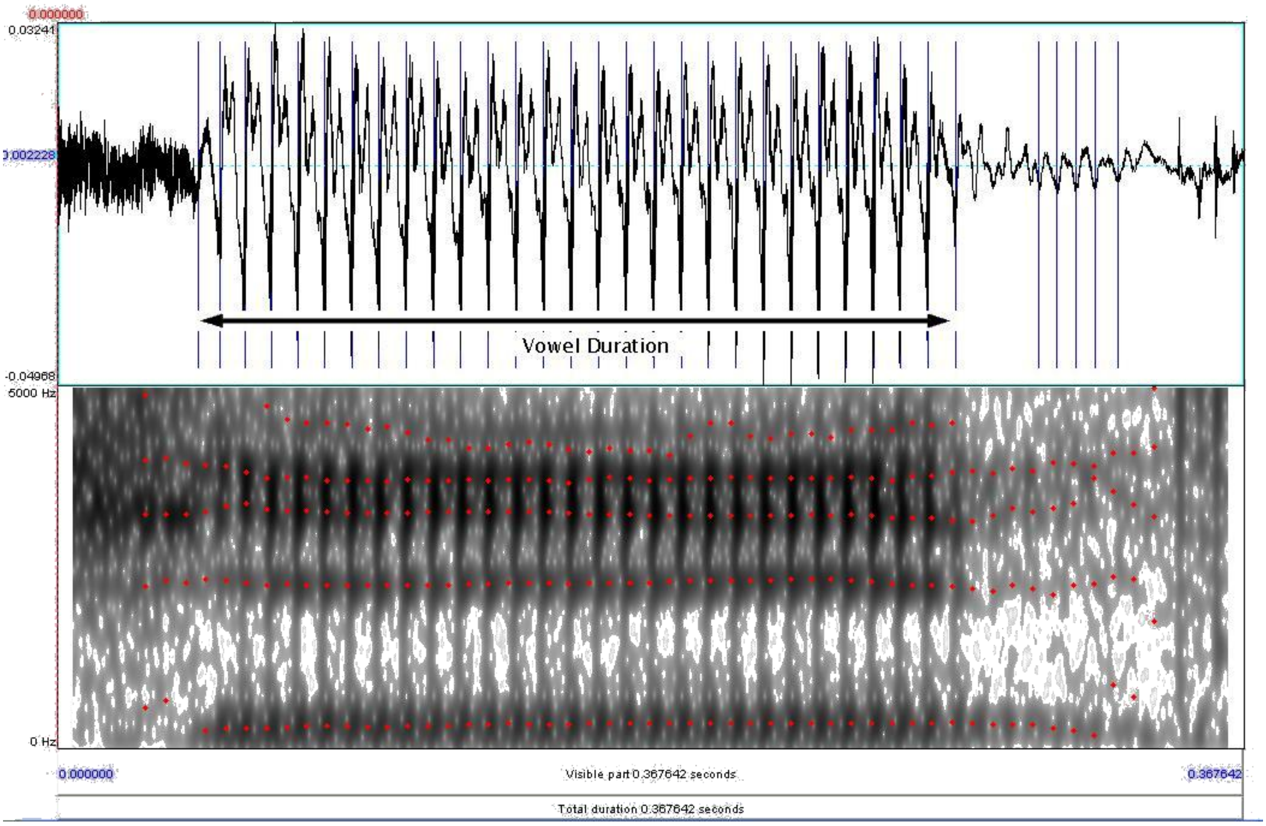


Figure 3. The waveform and wide-band spectrogram of the vowel /i/ produced in the context of "heed".

Measurements of vowel duration and formant frequencies were similar to those reported by Chen et al. (2001). Two vertical cursors were placed on the waveform to indicate the vowel duration: the first one was placed at the onset of the glottal cycle of vowel portion;

and the other one was at the offset of the glottal cycle. The offset of each vowel was evidenced by the abrupt cessation of energy in the lower formant (Port & Rotunno, 1979). The vowel duration was indicated by the time between the two cursors. Formant frequency measurements were calculated from the medial 80% of the vowel. The wide-band spectrogram obtained by using FFT was used to evaluate F1 andF2 values. Formant tracings using the built-in LPC algorithm were superimposed on the spectrogram. F1 and F2 values were calculated frame by frame. These values were averaged to obtain the mean F1 and F2 values.

In order to differentiate between tense and lax vowels, differential duration (DD) and average Euclidean distance (ED) between the tense and lax vowels were calculated. Differential duration (DD) is defined as the difference between the vowel duration of tense and lax vowels: D_{tense}-$D_{lax.}$ A smaller DD indicates a smaller tense/lax differentiation in duration of the vowels. The calculation of average Euclidean distance (ED) was given by:

$$ED = \sqrt{\left[\left(\mathrm{F1}_{\mathrm{lax}} - \mathrm{F1}_{\mathrm{tense}}\right)^2 + \left(\mathrm{F2}_{\mathrm{lax}} - \mathrm{F2}_{\mathrm{tense}}\right)^2\right]}$$

It was based on the assumption that F1 and F2 values were presented in a two-dimensional vowel space (Chen, 2006). ED indicates the distance between the F1-F2 location of tense and lax vowels distance in vowel space. It reflects the amount of separation between the two vowels.

Mean and standard deviation values of F1, F2, and duration of each vowel produced by Cantonese speaker were calculated by using SPSS 14.0. Paired sample t- tests were performed to determine possible differences in DD and ED between the vowels in each pair (/i-ɪ/, /æ -ɛ/, /u-ʊ/). Data indicate whether there is significant difference in acoustic properties between vowels in each pair. A 2-way ANOVA was performed to determine if significantly difference are present between gender and duration for each vowel pairs.

DD

Mean and standard deviation values of DD produced by Cantonese speakers were calculated. Based on the normative data of duration of each vowel obtained from American English, average DD of each vowel pair produced by American English speakers were also calculated. Descriptive comparison DD for each vowel pairs group differs between Cantonese speaker of English and American English speakers was provided. It is predicted that there will be great difference between all vowel pairs in DD with the Cantonese speakers producing shorter DD because tense/lax vowel is not present in the Cantonese vowel system.

ED

Average ED values for each vowel pair produced by Cantonese speakers were calculated by using the average the F1 and F2 values of each vowel according to the equation previously discussed. Average ED values for each vowel pairs produced by native English speakers were also calculated using data reported by Hillenbrand, Getty, Clark, and Wheeler (1995). To evaluate whether ED for each vowel pairs group differs between Cantonese speaker of English and native English speakers, descriptive analysis was used. It is predicted that there will also be great difference between groups in all vowel pair, the Cantonese speaker of English will show shorter ED because Cantonese speaker will not perceive and thus produce the contrast with difficulty if there is not such contrast in Cantonese (L1). The difference in ED between Cantonese speaker and native English speaker will be greater than that of DD because Cantonese vowel system has vowel allophones which differ in duration of time. It is easier for them to perceive difference in duration rather than vowel space.

Perceptual analysis

A group of 10 naïve listeners who were native speakers of American English was recruited for the perceptual analysis. Prior to the analysis, one out of three trials of each tense/lax vowel pair was chosen randomly. Each Cantonese speakers' production of tense/lax vowel in the /h**V**d/ context was separated and presented to the listeners randomly using a Microsoft PowerPoint file (3 vowel pairs x 40 speakers x 1 trial = 120 slides) (Appendix B). A practice period of about 10 minutes was given to the listeners in order to familiarize them with the procedures. The vowels were presented in pairs: /i-ɪ/, /æ -ɛ/, /u-ʊ/ inserted in each slide. The listening task was carried out in a quiet room. Each sample pair was presented to the listener. Listeners were asked to judge and determine whether vowel pair perceived was "different", "not different" or "not sure". An answer sheet was provided to the listeners (Appendix C). Each vowel pair was allowed to repeat once. Intelligibility index for each tense/lax pair was calculated as ratio (%) of the total number of vowel pairs perceived as "different", "not different", and "not sure" to the total number of vowel pairs (i.e. 60 male pairs and 60 female pairs). Comparison was made between the vowel groups and between the productions made by Cantonese and native English speakers. It is predicted the intelligibility index for each tense/lax vowel pair group will be lower for Cantonese production than that of Native English production.

Acoustic Measurement Reliability

Ten percent of all the recordings of /hVd/ vowel contexts (72 vowels) were randomly selected for testing the intra-rater and inter-rater reliability. The researcher extracted the F1, F2, and the duration from the 72 vowel contexts again for the intra-rater reliability. Another individual studying speech and hearing sciences performed the inter-rater reliability. Average absolute error for vowel duration intra-rater reliability was 8.3 ms and inter-rater reliability was 8.5 ms. The Pearson correlation coefficient for vowel duration intra-rater reliability was 0.96 ($p<0.01$) and intra-rater reliability was 0.95 ($p<0.01$). Average absolute error for F1 (in Hz) intra-rater reliability was 41.9 Hz and inter-rater reliability was 44.6 Hz. The Pearson correlation coefficients for F1 intra-rater reliability was 0.91 ($p<0.01$) and for F1 inter-rater reliability was 0.91 ($p<0.01$). Average absolute error for F2 (in Hz) intra-rater reliability was 48.8 Hz and inter-rater reliability was 50.6 Hz. The Pearson correlation coefficients for F2 intra-rater reliability was 0.98 ($p<0.01$) and for F2 inter-rater reliability was 0.98 ($p<0.01$).

Table 1a. *Mean (M), standard deviation (SD), and range values in Hertz (Hz) for the F1 and F2 frequency and in millisecond (ms) for duration of 3 tense-lax vowel pairs (/i-ɪ/, /æ -ɛ/, /u-ʊ/) produced by Cantonese male speakers.*

Vowel	N	F1 (Hz)			F2 (Hz)			Duration(ms)		
		Range	M	SD	Range	M	SD	Range	M	SD
/æ/	60	481 -980	648	109	1537 -2256	1816	155	79 -246	166	39
/ɛ/	60	412 -803	582	92	1733 -2352	1939	147	73 -262	151	42
/i/	60	242 -381	305	33	2094 -2914	2367	178	114 -253	168	38
/ɪ/	60	255 -445	343	43	1922 -2681	2252	181	55 -169	102	28
/u/	53	257 -414	348	32	503 -1861	896	248	107 -309	200	54
/ʊ/	53	275 -447	354	43	625 -1703	980	242	90 -263	162	39

Table 1b. *Mean (M), standard deviation (SD), and range values in Hertz (Hz) for the F1 and F2 frequency and in millisecond (ms) for duration of 3 tense-lax vowel pairs (/i-ɪ/, /æ -ɛ/, /u-ʊ/) produced by Cantonese female speakers.*

Vowel	N	F1 (Hz)			F2 (Hz)			Duration(ms)		
		Range	M	SD	Range	M	SD	Range	M	SD
/æ/	60	628 -1123	795	110	1625 -2920	1953	195	79 -246	166	39
/ɛ/	60	558 -912	739	79	1759 -2447	2044	154	112 -250	165	34
/i/	60	270 -870	379	92	2016 -3068	2766	208	86 -292	176	43
/ɪ/	60	286 -624	436	72	1957 -2960	2505	280	59 -165	109	24
/u/	59	354 -514	411	34	751 -1509	1008	178	123 -302	194	39
/ʊ/	59	306 -531	416	43	861 -1677	1125	193	96 -274	168	46

Results

Acoustic analysis

To investigate the acoustic and temporal characteristics of tense and lax vowels produced by Cantonese speakers of English, F1 and F2 values, and duration of each vowel were determined. Mean and standard deviations of F1, F2 values, and duration for the three tense/lax vowel pairs (/i-ɪ/, /æ -Ɛ/, /u-ʊ/) produced by male and female Cantonese speakers of English are depicted in Tables 1a and 1b. Two-tailed paired-sample t-tests were carried out to determine if F1, F2 and duration for each tense vowel were significantly different from the corresponding lax vowels.

Results of statistical testing indicated that duration, F1 and F2 frequency values associated with the tense vowel /i/ were significantly different from those with the lax vowel /ɪ/ produced by Cantonese speakers. For vowel pairs /æ–Ɛ/ and /u-ʊ/, Cantonese speakers produced significantly different duration and F2 in vowel pair /u-ʊ/ and significantly different F1 in vowel pair /æ–Ɛ/. There were no difference between /æ–Ɛ/ vowel pairs in duration and F2 and /u-ʊ/ vowel pairs in F1.

A two-way Analysis of Variance (ANOVA) was carried out in order to determine whether there was any interactive effect between gender and duration of vowels. Results showed that significant interaction effects (Gender x Duration) only existed in /æ–Ɛ/ while there is no significant difference in duration produced by female and male Cantonese speakers of English. Figures 1a and 1b show the interaction (Gender x Duration) in different tense/lax vowel pairs between gender and duration.

Table 2. *Mean (M), standard deviation (SD), and range of the DD (ms) for 3 pairs tense-lax vowels /i-ɪ/, /æ -Ɛ/, /u-ʊ/ produced by Cantonese male and female speakers.*

		N	Range	Mean±SD
Male	/i-ɪ/	60	-36 -176	66 ±43
	/u-ʊ/	53	-104 -179	38 ±69
	/æ–Ɛ/	60	-165 -118	15 ±60
Female	/i-ɪ/	60	-27 -199	67 ±53
	/u-ʊ/	59	-74 -196	26 ±58
	/æ–Ɛ/	60	-111 -113	1 ±44

Mean and standard deviation values of DD (ms) of the three tense-lax vowel pairs (/i-ɪ/, /æ -Ɛ/, /u-ʊ/) produced by the Cantonese speakers are depicted in Table 2. The results were compared with the normative data of American English speakers (see Figure 4).

Results indicated that /u-ʊ/ and /æ–Ɛ/ were produced with relatively shorter DD whereas /i-ɪ/ exhibited comparable duration with that produced by native American English speakers. The greatest difference in DD between Cantonese and American was observed in the tense-lax vowel contrast of /æ–Ɛ/. The order of DD difference between Cantonese and American is: /i-ɪ/</u-ʊ/<./æ–Ɛ/.

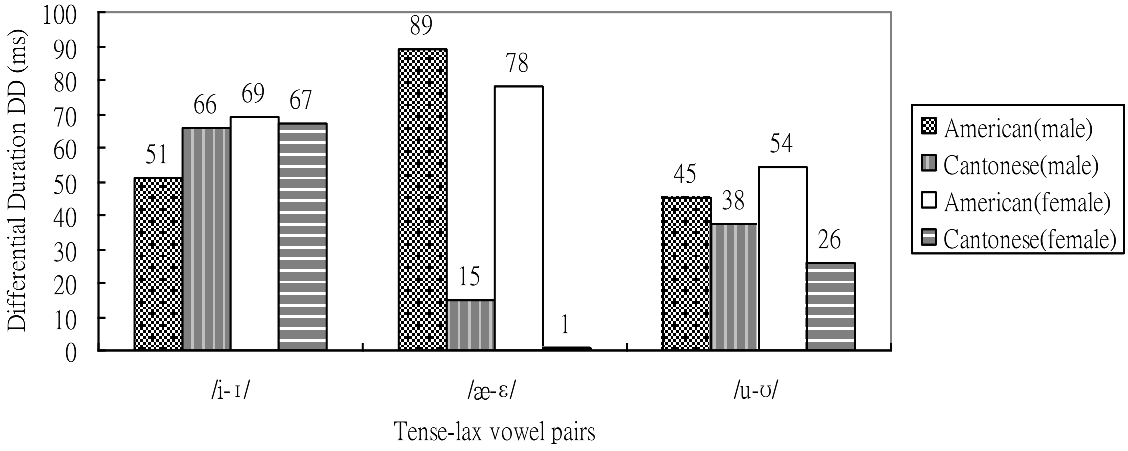


Figure4a. Comparison of Differential Duration (DD) of the three pairs tense-lax contrast /i-ɪ/, /æ -ɛ/, /u-ʊ/ between American English speakers and Cantonese speakers of English.

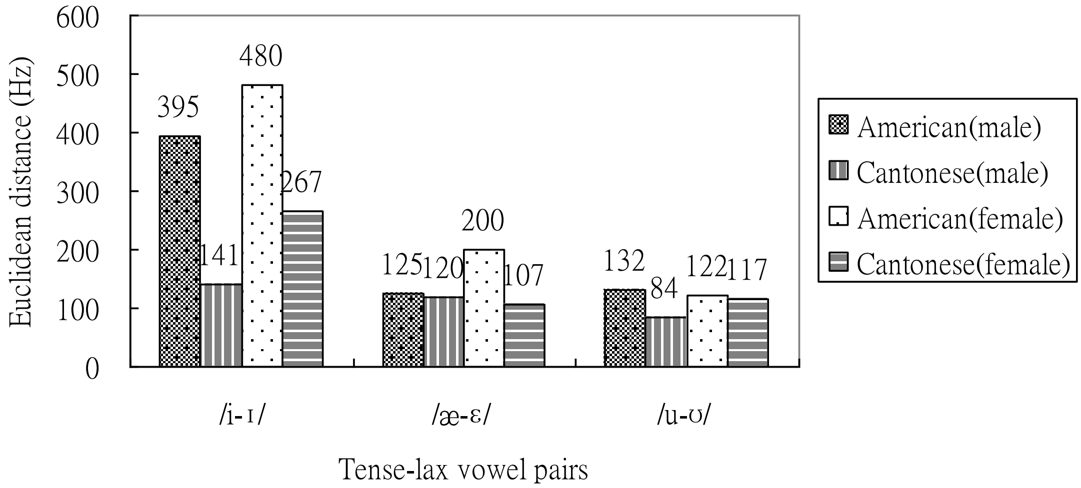


Figure4b. Comparison of Euclidean distance (ED) of the three pairs tense-lax contrast /i-ɪ/, /æ -ɛ/, /u-ʊ/ between American English speakers and Cantonese speakers of English.

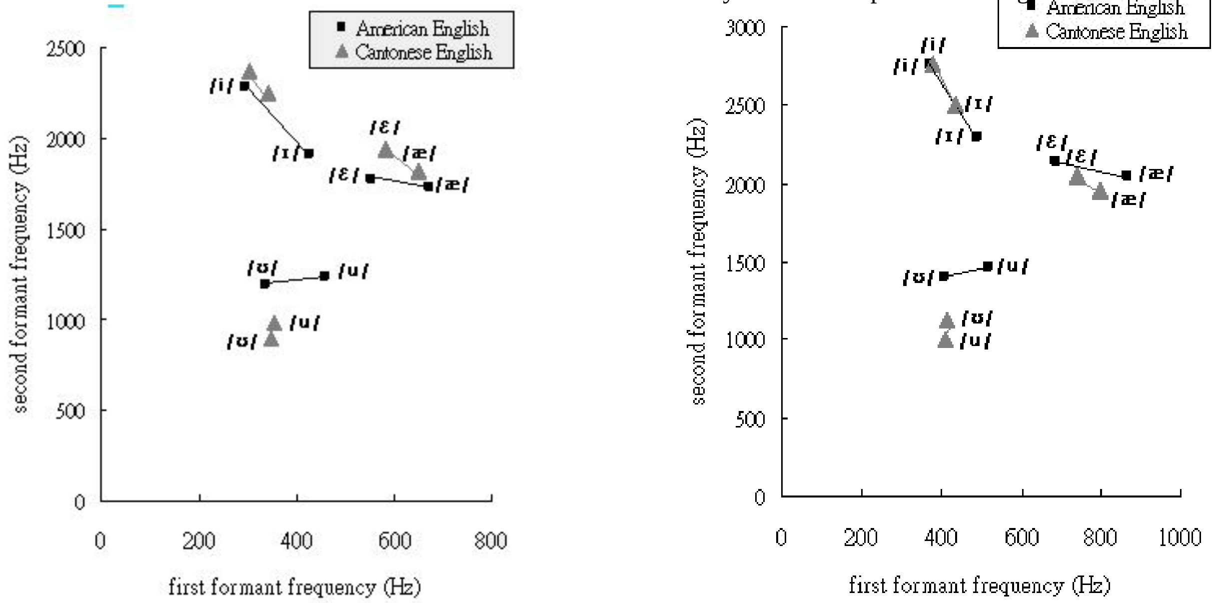


Figure 4c. Comparison of ED of the tense-lax contrast in an F1x F2 plot between male (left) and female (right) speakers of Cantonese and American.

Table 3. *Average ED values (Hz) for the tense-lax vowel pairs (/i-ɪ/, /æ -ɛ/, /u-ʊ)/ produced by Cantonese female speakers.*

		N	Mean
Male	/i-ɪ/	60	141
	/u-ʊ/	53	84
	/æ–ɛ/	60	120
Female	/i-ɪ/	60	267
	/u-ʊ/	59	117
	/æ–ɛ/	60	107

Table 4. *Percentage of tense-lax vowel pairings produced by Cantonese speakers of English that were perceived by Native Americans as distinctly different (n=10)*

	Males			Females		
	Different	Not different	Not sure	Different	Not different	Not sure
/i, ɪ/	43.18%	43.18%	13.64%	50.00%	43.18%	6.82%
/ε, æ/	45.46%	50.00%	4.55%	40.91%	47.73%	11.36%
/u, ʊ/	20.46%	77.27%	2.27%	15.91%	75.00%	9.09%

The average ED values produced by Cantonese speakers were calculated and summarized in Table 3. The difference in ED values between Cantonese speakers and American English speakers are shown in Figures 4b & 4c. The ED differences between productions made by American and Cantonese speakers indicated that Cantonese speakers tended to produce shorter Euclidean distance (ED) in all the three vowel pairs /i-ɪ/, /æ -ɛ/, /u-ʊ/. For both genders, the longest ED was found in /i-ɪ/ vowels compared with the other two pairs. The order of ED values among the three pairs is: /i-ɪ/>/æ -ɛ/>/u-ʊ/.

Perceptual Analysis

The average intelligibility scores for the tense-lax English vowel pairs produced by Cantonese speakers are shown in Table 4. Results indicated that female and male adults who were native speakers of English showed greatest difficulty in differentiating between /u/ and /ʊ/ produced by Cantonese speakers. Over 70% of the /u-ʊ/ were perceived as “not different”. The order of intelligibility scores of the vowel pairs is: /i-ɪ/>/æ–ɛ/>/u-ʊ/.

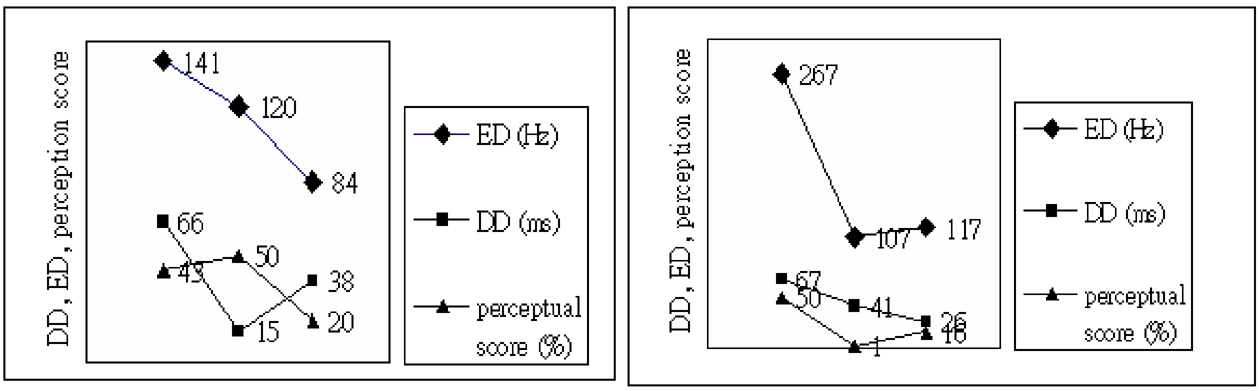
Discussion

Data from the analyses indicated that Cantonese speakers of English only produced significantly different /i/ and /ɪ/ for all the three parameters (F1, F2, and duration), whereas F1 was not significantly different between /u/ and /ʊ/, and /æ/ and /ɛ/ were produced without significantly difference in F2 value and duration. Since vowel categorization depends on the first two formant frequency in the acoustic plane, it is difficult to distinguish between /u/ and /ʊ/, and /æ/ and/ɛ/ if the two vowels within the pair do not have significant different in either F1 or F2, (Rosner & Pickering, 1994).

The present acoustic results appear to be consistent with the findings from perceptual tasks which showed that /i-ɪ/ was perceived as “distinctively different” with comparatively highest overall percentage. As ED reflects the “acoustic distance” between the vowels, the greater the DD and ED between /i/ and /ɪ/ imply that the vowels should be more distinguishable. It was observed that DD of /i-ɪ/ produced by Cantonese speakers was comparable with that by native American English speakers, and it was the longest among the three vowel pairs. Euclidean distance (ED) of /i-ɪ/ was also the greatest among the three vowel pairs though it was smaller than that of American English. The longer DD and greater ED provided salient temporal and spectral cues for the listeners to distinguish between /i/ and /ɪ/ effectively.

The perceptual scores obtained from /i-ɪ/ contrast are in agreement with the study reported by Chen (2006) who found out that vowel produced by Mandarin was perceived as “different” with the percentage around 50%. It was suggested in previous studies that more language exposure facilitates language learning. If the vowel occurs frequently in the target language, it will facilitate the acquisition of accurate vowel sound articulation (Chen, 2006). Study of English words has found that relative percentage of occurrence frequency for /i-ɪ/, /æ−Ɛ/ and /u-ʊ/: 8.88%, 5.7%, and 2.25%, respectively (Chen, 2006). The highest occurrence frequency of /i-ɪ/ in English increase Cantonese speakers' language experience to be equipped with more accurate temporal and acoustic properties in articulating /i-ɪ/ (Rosner & Pickering, 1994).

However, the acoustic results contradicted with data from perceptual evaluation. Though /æ/ and /Ɛ/ were not significantly different in duration and F2, vowel contrast was perceived with higher percentage of "distinctively different" than that of vowel pair /u-ʊ/ which has significantly different duration and F2. This may be due to the relatively longer ED values of /æ−Ɛ/ than /u-ʊ/ (ED of /æ−Ɛ/ was 36 ms longer than that of /u-ʊ/ for male speakers, though 10 ms shorter for female speakers)(see figure 5). Such longer ED values indicate that there is a great change in vowel space which determines the differentiation of tense lax vowel (Chen, 2006).



/i-ɪ/ /æ−Ɛ/ /u-ʊ/ **/i-ɪ/ /æ−Ɛ/ /u-ʊ/**

Tense lax vowel pairs Tense lax vowel pairs

Figure 5. Comparison of ED, DD, and perceptual score for the three tense lax vowel pairs

/i-ɪ/, /æ–ɛ/ and /u-ʊ/ produced by Cantonese male (left) and female (right).

Furthermore, /æ/ and /ɛ/ were produced with tonal difference as "had" and "head" by Cantonese speakers. The word "head" tends to be produced with a rising tone by Cantonese speakers of English, while "had" pronounced with a falling intonation. Data from Cantonese male speakers produced the word "head" with fundamental frequency (F0) (average F0 = 219 Hz) 22 Hz greater than that of "had" (average F0 = 196 Hz). Cantonese female speakers produced the word "head" with average F0 of 131 Hz which is 12 Hz greater than that of "had" (average F0 = 119 Hz). Vowel quality can be changed by subtle factors, like stress level of the syllables, neighboring consonants, etc. (Olive et al., 1993). This subtle difference in tone of "had" and "head" may be salient to the listeners which allow them to perceive as "different". The phenomenon can be observed from Figure 3 in which, with decreasing ED and DD values, the perceptual score of /æ/ and /ɛ/ remains relative high (around 50%) while the perceptual score keeps falling with the decrease in ED in /u-ʊ/.

Data from the perceptual and the acoustic experiments indicate that Cantonese speakers made use of durational cues as an important feature for producing tense/lax contrast. The vowels /i/ and /ɪ/ were produced with a DD value that was comparable with that of American English, and it was over 30-50 ms longer than those between /æ/ and /ɛ/, and /u/ and /ʊ/. At the same time, the /i-ɪ/ vowel pair was differentiated with the highest perceptual score, despite the large discrepancy between the ED of /i-ɪ/ produced by Cantonese and American. On the contrary, as /æ–ɛ/ and /u-ʊ/ were produced with relatively short DD, the percentage of “not different” responses increased from 43% in /i-ɪ/ to 50% in /æ–ɛ/ and 77% in /u-ʊ/ for male and from 43% in /i-ɪ/ to 48% in /æ–ɛ/ and 75% in /u-ʊ/ for male. This indicates that temporal cue is not the primary factor in differentiating between tense and lax vowels. Spectral cue appears to weigh more in distinguishing the vowels. The greater ED provides greater contrast between tense and lax vowels, and therefore, /æ–ɛ/ were more distinguishable and yielded a higher percentage of “distinctively different” responses. This rule of tense-lax vowel differentiation was also observed in Mandarin Accented English (MAE) as reported by Chen (2006). He found that MAE relied on temporal difference to maintain vowel contrast and American English uses primarily spectral contrast to determine tense lax vowels.

The English vowel pairs produced by native Cantonese speakers were less distinguishable than those produced by American English speakers. This can be shown by the relatively shorter ED and DD values in /i-ɪ/, /æ–ɛ/ and /u-ʊ/ produced by Cantonese speakers. Cantonese speakers might acquire sufficient skills in integrating the temporal and spectral properties for a more accurate articulation of tense/lax vowel pairs. These lead to invariant acoustic features in Cantonese English The different acoustical properties in the

English tense and lax vowels produced by American English speakers and Cantonese speakers can be explained by the Contrastive Analysis Hypothesis: phonological features of L1 (Cantonese) that are similar to that of L2 (English) are transferred or facilitate the acquisition of L2 (English) (Flege, et al, 1997; Jenkins, 2000). As there is no tense/ lax contrast in Cantonese, this may lead to difficulties in perceiving the spectral features in order to differentiate the tense/ lax contrast produced by English speakers. They fail to manipulate the vowel space (F1 and F2 values) to yield distinctive tense and lax vowels. Articulation of tense and lax vowels cannot be shifted from a more extreme position in tense vowel to a more centralized location in lax vowel (Ladefoged, 1975).

Apart from the innate perception ability of Cantonese that affect the acquisition of correct tense/lax vowels, the learning environment may also be another factor affecting the acquisition of spectral contrast of tense and lax vowels in the Cantonese-speaking community in Hong Kong. In Hong Kong, English is taught as a second language by teachers who are also native Cantonese speakers who also lack the ability to perceive/produce the spectral contrast between tense and lax vowels. The teachers may make use of temporal cues to differentiate the vowels. Language experience in tense and lax vowels with insufficient spectral contrast but salient temporal cues equip the Cantonese speakers to rely primarily on the temporal feature to maintain the tense/lax vowel difference.

Furthermore, since there are long and short allophones in the Cantonese vowel system, Cantonese speakers may transfer the ability in managing the temporal characteristics of long and short allophones to tense and lax vowels. They may articulate the English tense and lax vowel pairs in a similar way as long and short allophones in Cantonese. Therefore, comparable DD or even higher DD in /i-ɪ/ but shorter ED were observed.

Results from the study of Mandarin (Chen, 2006) and the current study revealed that a greater increase in DD in tense/lax vowel pairs produced by Cantonese compared to those produced by Mandarin. For example, the average DD value of /u-ʊ/ was reported to be 8 ms in Mandarin (Chen, 2006), but the current study shows that Cantonese produced an average DD of 26 ms for females and 38 ms for males. The presence of allophones may account for such phenomenon. Mainly different in duration, the allophones present in the Cantonese sound system familiarize the native speakers of Cantonese with the duration changes in production of vowels (Bauer & Benedict, 1997). The Cantonese speakers may be more capable of managing the temporal properties of the tense/lax vowels.

Implications and Future Research

The findings from the present study can help understanding the acquisition of English as a second language and support the Contrastive Analysis Hypothesis. However, since the study only made use of normative data of American English reported previously, statistical tests of Euclidean distance and Differential Distance between American English and Cantonese English are not available. Future studies should include native speakers of American English for production and perception tasks. The notion of limited acoustic diversity in native Cantonese speakers producing tense and lax English vowels can then be tested statically.

Further investigation on how allophones in Cantonese influence the acquisition of tense and lax English vowels in comparison to Mandarin speakers can also be carried out. Possible differences in DD and ED values between Cantonese and Mandarin speakers can be determined. If the DD value associated with Cantonese speakers are significantly longer than Mandarin speakers, it can provide support to the transfer of temporal modification from allophones production to tense and lax vowels.

References

Bauer, R. S. & Benedict, P. K. (1997). *Modern Cantonese phonology.* Berlin: Mouten de Cruyter.

Chen,Y. (2006). Production of tense-lax contrast by mandarin speakers of English. *Folia Phoniatrica et Logopaedica, 58*, 240-249.

Chen, Y., Robb, M., Gilbert, H., & Lerman, J. (2001). Vowel production of Mandarin speakers of English. *Clinical Linguistics & Phonetics*, 15, 427-440.

Colton, R. H. & Casper, J. K. (1996). *Understanding voice problems: a physiological perspective for diagnosis and treatment.* Baltimore: Williams & Wilkins.

Eckman, F. R., Elreyes, A., & Iverson, G. K. (2003). Some principles of second language phonology. *Second Language Research, 19,* 3. pp. 169-208.

Edwards, H. T. (2003). *Applied Phonetics: The sounds of American English. (3rd ed).*

Flege, J. (1989). Chinese subjects' perception of the word-final English /t/-/d/ contrast performance before and after training. *Journal of the Acoustical Society of America, 86,* 1684-1697.

Flege, J., Bohn, O., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception. *Journal of Phonetics, 25,* 437-470.

Hillenbrand, J., Getty, L. A., Clark, M. J., & Wheeler, K. (1995). Acoustic characteristics of American English Vowels. *Acoustical Society of America, Vol. 97, Pt 1.*

Hung, T. T. N. (2000). Towards a phonology of Hong Kong English. *World Englishes, 19,* 337-356.

Jenkins, J. (2000). *The phonology of English as an international language.* Oxford : Oxford University Press. Pp.99-119.

Kewley-Port, D. & Atal, B. (1989). Perceptual differences between vowels located in a limited phonetic space. *Journal of the Acoustic society of America, 85*, 1726-1740.

McAllister, R., Flege, J. E., & Piske, T. (2002). The influence of L1 on the acquisition of Swedish quantity by native speakers of Spanish, English and Estonian. *Journal of Phonetics, 30*, 229-258.

Olive, J. P., Greenwood, A., & Coleman, J. (1993). *Acoustics of American English Speech: A dynamic speech.* New York: Springer-Verlag.

Ladefoged, P. (1975). *A course in phonetics.* New York: Harcourt Brace Jovanovich.

Wong, P. C. M., & Diehl, R. L (2003). Perceptual normalization for inter and intratalker variation in Cantonese level tones. *Journal of Speech, Language and Hearing Research*, 46, 2.

Peng, L. & Setter, J. (2000). The emergence of systematicity in the English pronunciations of two Cantonese-speaking adults in Hong Kong. *English World-Wide, 21*, 1.

Peterson, G. E., & Barney, H. E. (1952). Control methods used in the study of vowels. *Journal of the Acoustical Society of America, 24*, 175-184.

Port, R., & Rotunno, R. (1979). Relation between voice-onset time and vowel duration. *Journal of the Acoustical Society of America, 97*, 3099-3111.

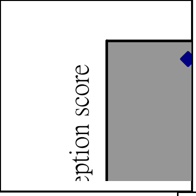
Rosner, B. S., & Pickering, J. B. (1994). *Vowel perception and production.* Oxford: Oxford University Press.

Appendix A

Appendix B

Appendix C

	√ difference	X difference	NOT sure		√ difference	X difference	NOT sure
1.				**54**			
2.				**55**			
3				**56**			
4				**57**			
5				**58**			
6				**59**			
7				**60**			
8				**61**			
9				**62**			
10				**63**			
11				**64**			
12				**65**			
13				**66**			
14							
15							
16							
17							
18							
19							
20							
21							
22							
23							
24							
25							
26							
27							
28							
29							
30							
31							
32							
33							
34							
35							
36							
37							
38							



39				
40				
41				
42				
43				
44				
45				
46				
47				
48				
49				
50				
51				
52				
53				

Acknowledgements

I are grateful to all the 40 subjects from the University of Hong Kong for their participation. I thank Manwa Lawrence Ng for his helpful advice and supervision. I also thank Mr. Raymond Wu for technical support.